Building “Cultural” Knowledge Fragments

Eugene Santos Jr.
Thayer School of Engineering
Dartmouth College
Eugene.Santos.Jr@dartmouth.edu
http://di2ag.thayer.dartmouth.edu

Distribution A: Approved for public release; distribution unlimited.

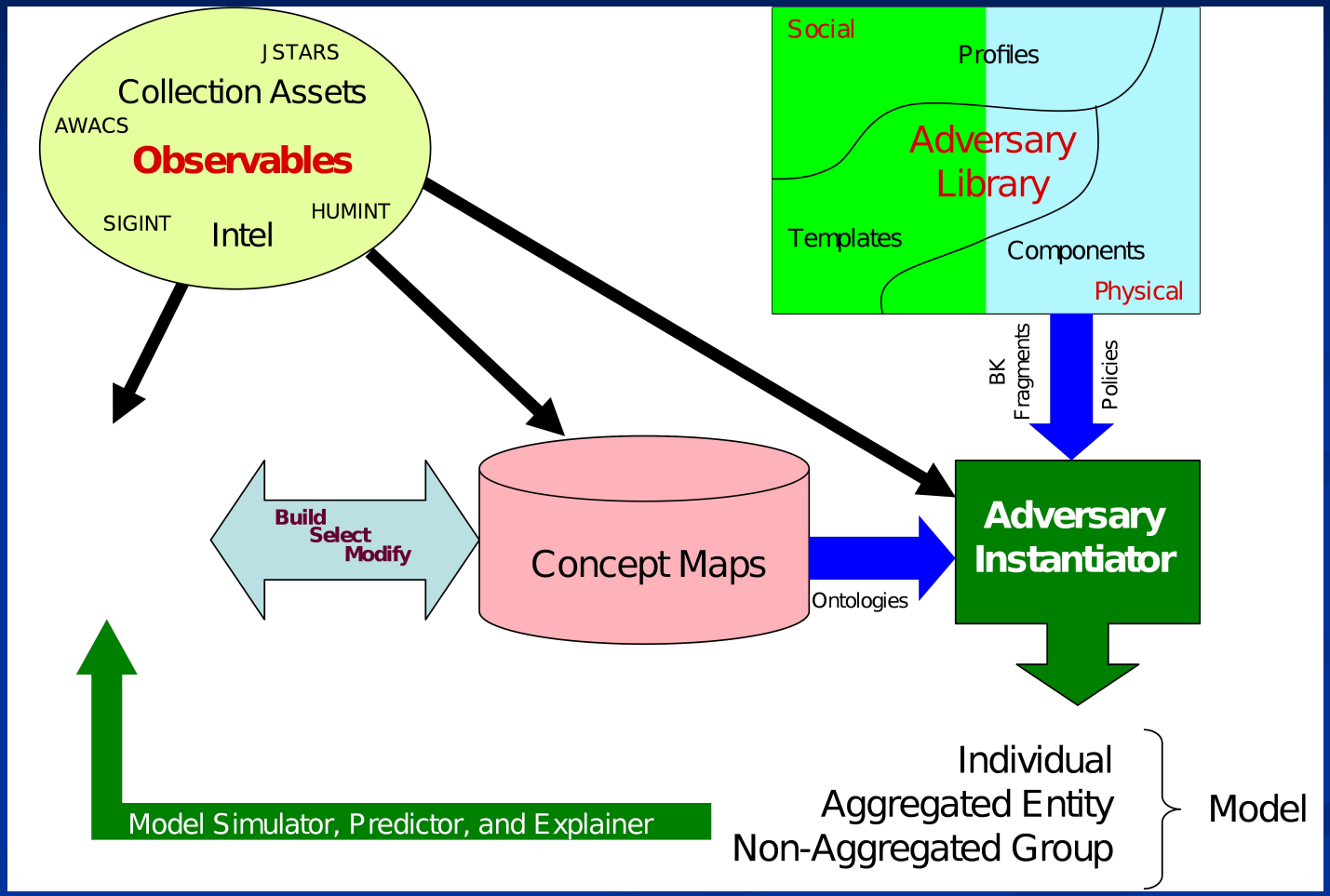
Team

- AFOSR Project: On the Effects of Culture and Society on Adversarial Attitudes and Behavior
  - Eugene Santos Jr. and Qunhua Zhao (Dartmouth) – computational adversarial modeling and Bayesian knowledge fragment library
  - Felicia Pratto (UConn) – cultural and social psychology of individuals and effects of groups
  - Jeff Bradshaw and Paul Feltovich (IHMC) – organizational behavior modeling and policy managements
  - Eunice E. Santos (Virginia Tech) – social networks analysis and computational testbeds
- Collaborations
  - Richard Warren (AFRL/HECS)
  - Duane Gilmour (AFRL/IFTC)
  - Lee Krause and Lynn Lehman (Securboration, Inc.)

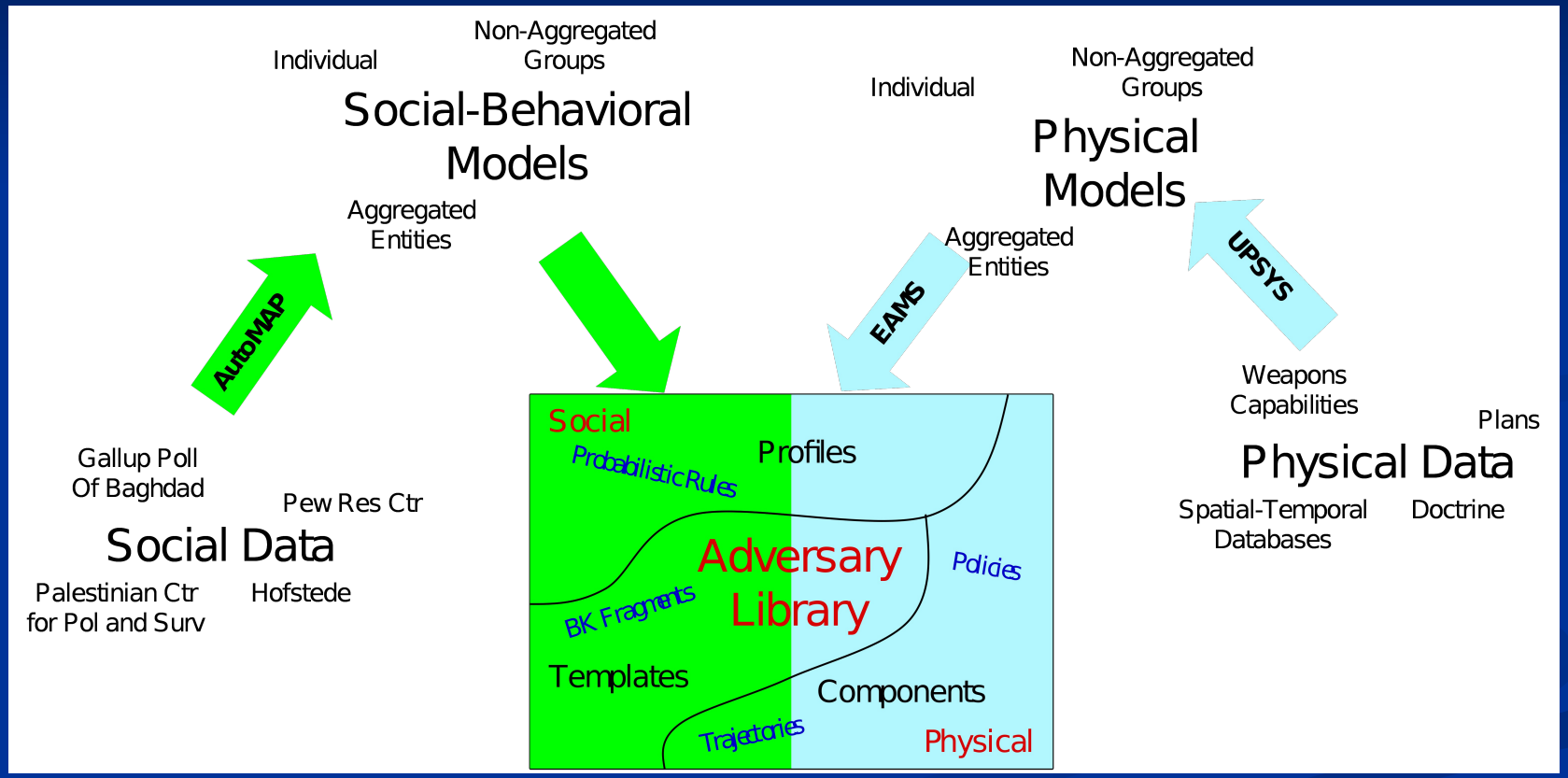
Objectives

- Design and develop a computational model for *inferring adversarial intent and predicting behavior*
- Build and employ social, cultural, and political data-driven models to *explore and explain* (in addition to modeling) adversarial attitudes and behaviors

Architecture



Adversary Library



What do you need to know about the adversary?

- Things like:
  - Histories of responses and actions in different situations?
  - Social/Economic/Military/Political/Religious doctrine?
  - Infrastructure and reliability of leadership or command and control?
  - Perceptions about us (our force) or other groups?
  - Political and cultural factors?
- Might provide clues on their propensity for future actions?
- What do we really need?

What is Intent?

- Intent inferencing, or user intent inferencing, involves deducing an entity's goals based on observations of that entity's actions (Geddes, 1986)
  - Deduction involves the construction of one or more behavioral models that have been optimized to the entity's behavior patterns
  - Data/knowledge representing observations of an entity, the entity's actions, or the entity's environment (collectively called observables) are collected and delivered to the model(s)
  - Models attempt to match observables against patterns of behavior and derive inferred intent from those patterns
- Userful for generation of advice, definition of future information requirements, proactive aiding, or a host of other benefits (Bell et al., 2002; Santos, 2003)

What is Adversary Intent?

- What's the context of a Red action?
- What is the rationale behind the Red action?
- What are the causes and effects of the intended Red goal?
- What is the motivation behind a Red behaviour?
- What will happen next?
- Why did this behaviour occur?
- What does Red believe?

Intent – What can you do with it?

- Predict the future: actions, reactions, behaviours, etc.
- Explain the present: causes, motivations, goals, etc.
- Understand the past: beliefs, axioms, history, etc.

- Inferred intent knowledge can help focus and prune search space, bound optimization, guide scheduling, and better allocate resources.

Adversary Intent

- Intent is not just the plan or enemy course of action
- Not just "The enemy commander *intends* to launch his SAMs" or "The organization *intends* to undertake a suicide bombing", but also why??

- Intent is the highest-level goal(s) the adversary is pursuing + the support for that goal + the plan to achieve it
- Need intent to understand and predict Red behavior
- Must model adversary based on their perceptions of the world

Focus of Talk

- "Cultural" knowledge fragments – human factors (elements) that define or influence decision-making central to a particular individual or organization
- Results thus far from modeling the intent behind suicide bombings in the middle east
  - Joint with Drs. Felicia Pratto and Qunhua Zhao

Accounting for Human Factors in Capturing Adversary's Intent

- Assymetric adversaries – they are not like us; we do not think like them
- "What is rational" is not the same between different individuals or groups especially with different backgrounds.
- Differences in decision-making and behavior come from differences in background
  - Social
  - Cultural
  - Economic
  - Political
  - Psychological

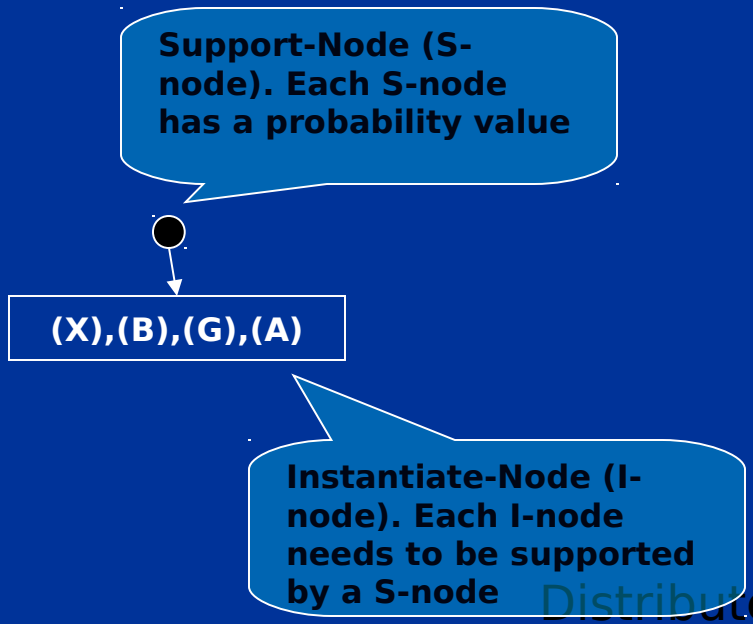
Challenges

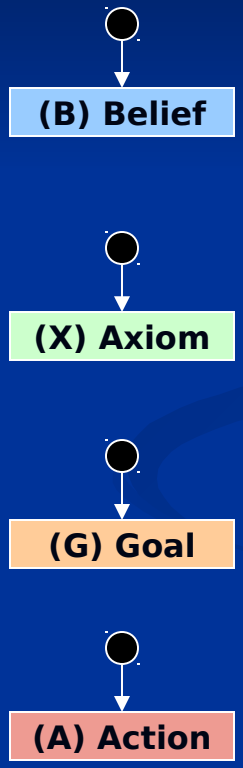
- Each individual or group is a unique entity
- Human factors are difficult to capture accurately and/or completely
- Uncertainty associated with the impacts of human factors on decision-making process is inherent

Our Adversary Modeling Approach

- Incorporate human factors
- Intent driven
- Model the decision making process based on how adversary views the world
- Build network fragments for each piece of information / knowledge, and merge them together for reasoning
- Based on Bayesian Knowledge Bases (BKBs)
- Fragments built and validated jointly with social scientist/subject matter experts

Basics for BKB fragments and Adversary Intent Inferencing Model


Support-Node (S-node). Each S-node has a probability value
(X),(B),(G),(A)
Instantiate-Node (I-node). Each I-node needs to be supported by a S-node



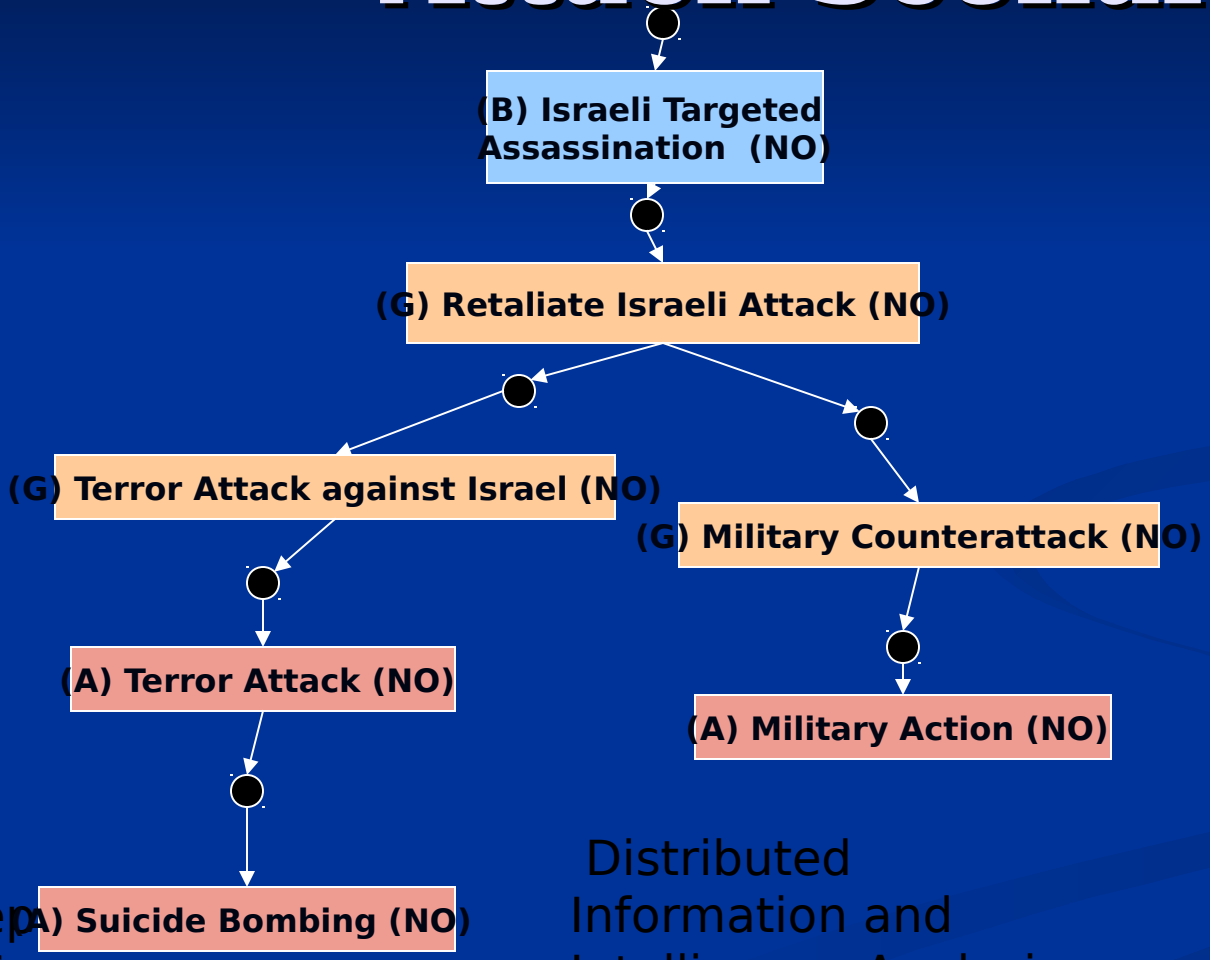
What the adversary believe about their opponents

What the adversary believe about themselves

What results the adversary wants to achieve

How they will carry out their tasks

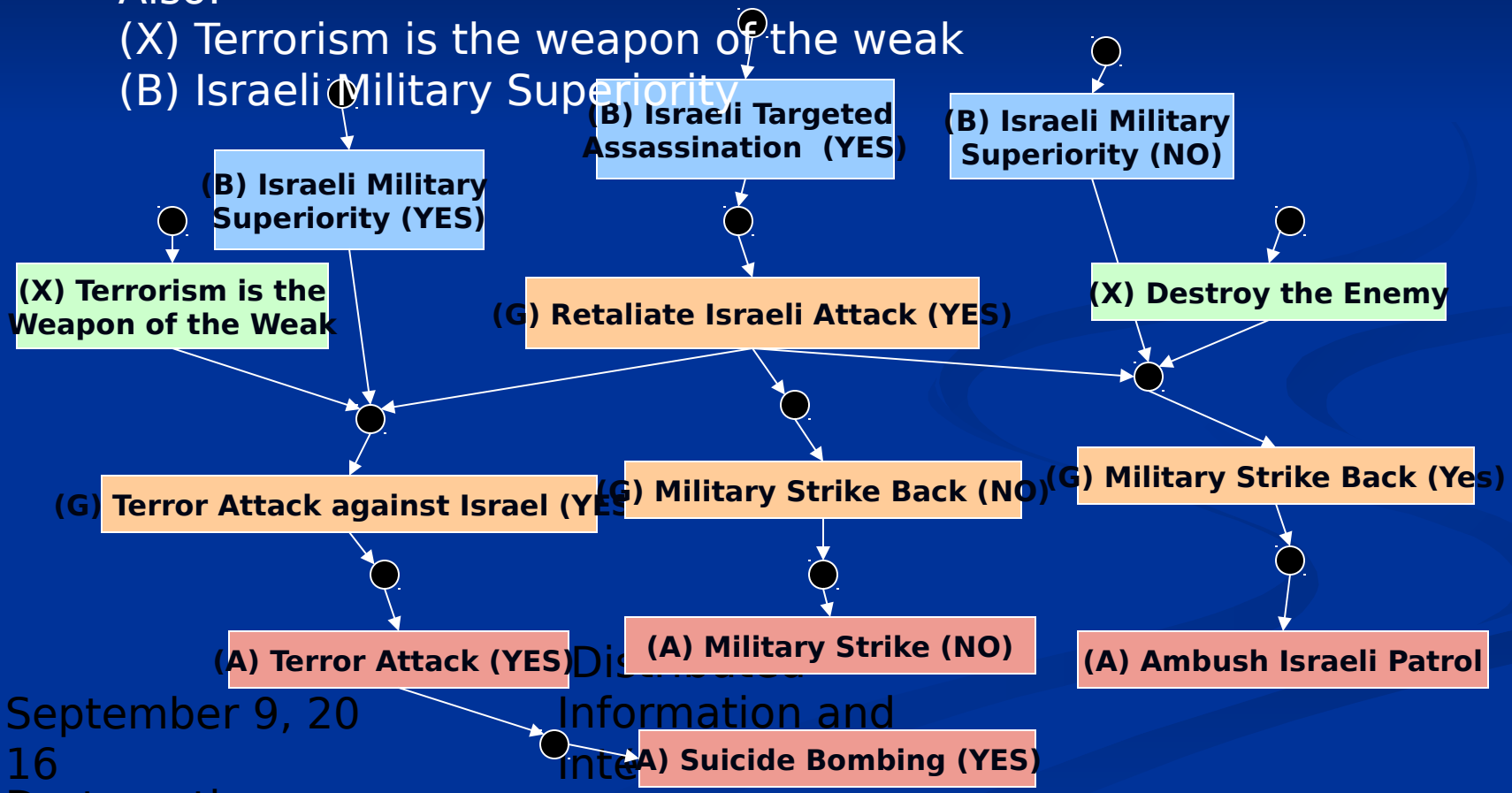
Constructing BKB Fragments from Terrorism Attack Scenario


(B) Israeli Targeted Assassination (NO)
(G) Retaliate Israeli Attack (NO)
(G) Terror Attack against Israel (NO)
(G) Military Counterattack (NO)
(A) Terror Attack (NO)
(A) Military Action (NO)
(A) Suicide Bombing (NO)

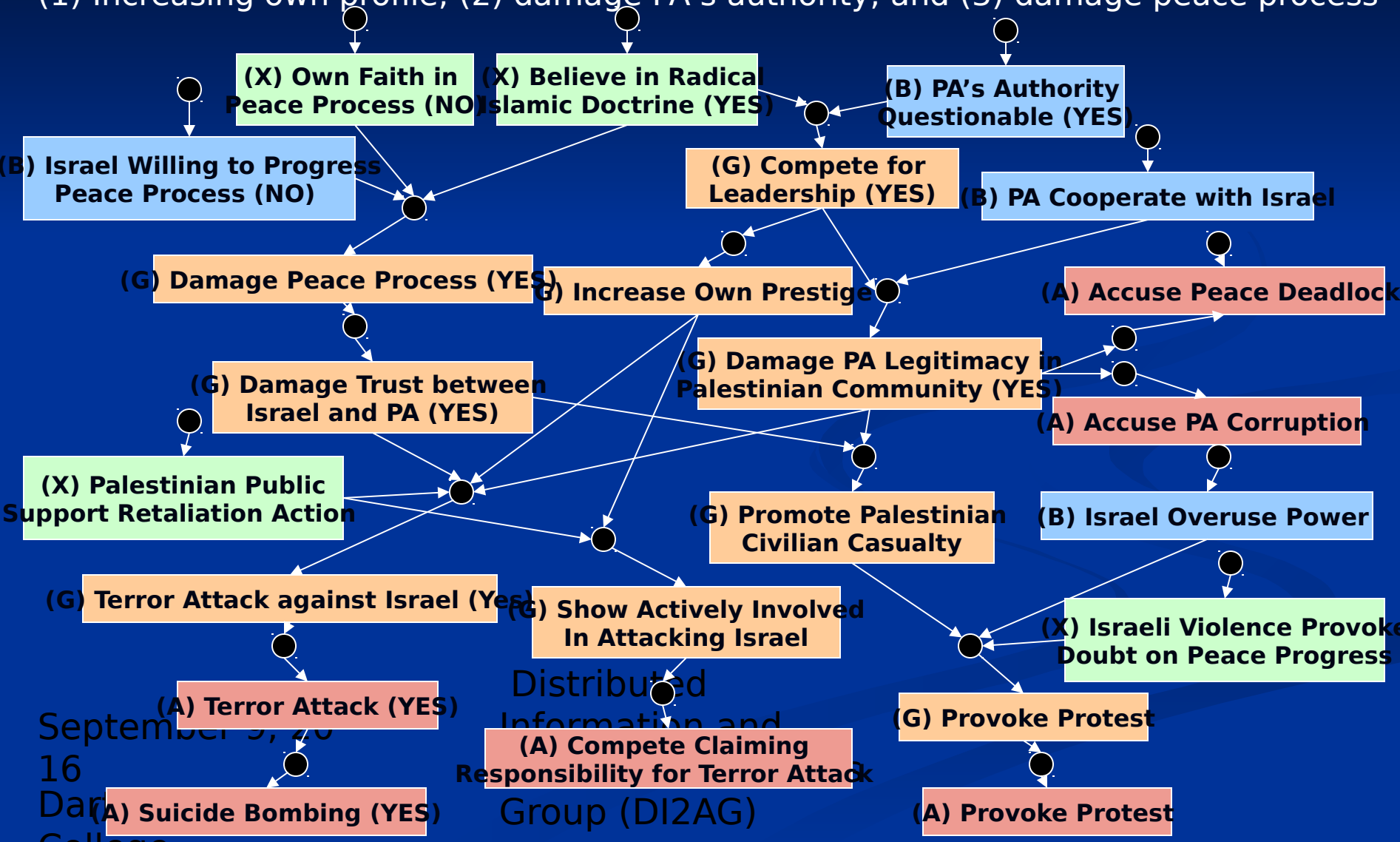
"Arafat convinced Hamas to suspend military actions after Sept. 11, 2001 on the condition that Israeli targeted assassination stop."

Mia Bloom (2005) "Dying to Kill, the allure of suicide terror"

An explanation follows from the logic that violence is often retaliatory;
“The al Ibrahimi Mosque massacre opened the doors of revenge in Palestinian like never before” (Mazin Hammad, cited in “Dying to Kill”).
Also:
(X) Terrorism is the weapon of the weak
(B) Israeli Military Superiority

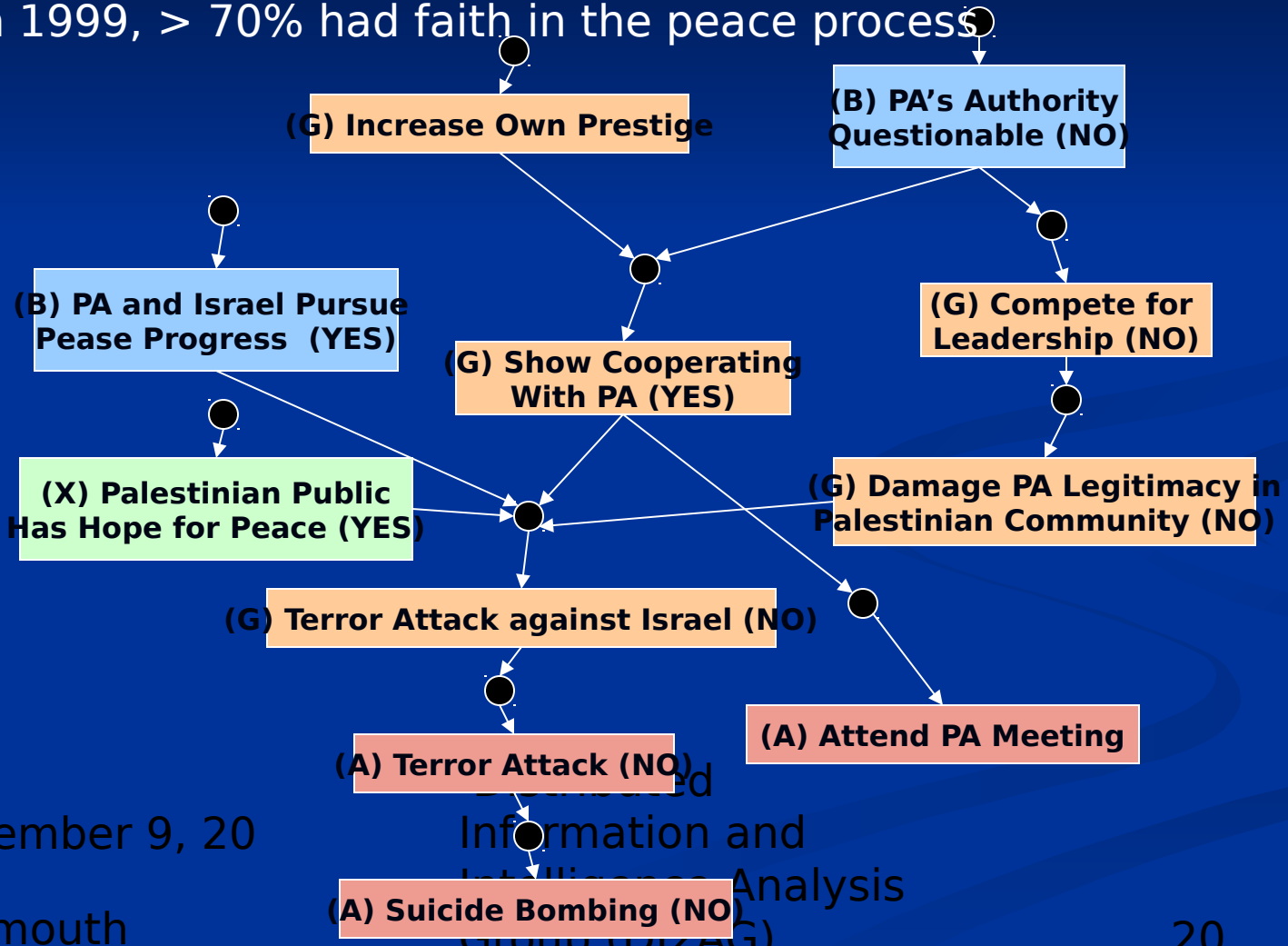

(B) Israeli Military Superiority (YES)
(B) Israeli Targeted Assassination (YES)
(B) Israeli Military Superiority (NO)
(X) Terrorism is the Weapon of the Weak
(G) Retaliate Israeli Attack (YES)
(X) Destroy the Enemy
(G) Terror Attack against Israel (YES)
(G) Military Strike Back (NO)
(G) Military Strike Back (Yes)
(A) Terror Attack (YES)
(A) Military Strike (NO)
(A) Ambush Israeli Patrol
(A) Suicide Bombing (YES)

Another view of the reason behind suicide bombing: Competing for the leadership in Palestinian community, when public has no hope in peace and supports violence for revenge.
(1) Increasing own profile; (2) damage PA's authority; and (3) damage peace process

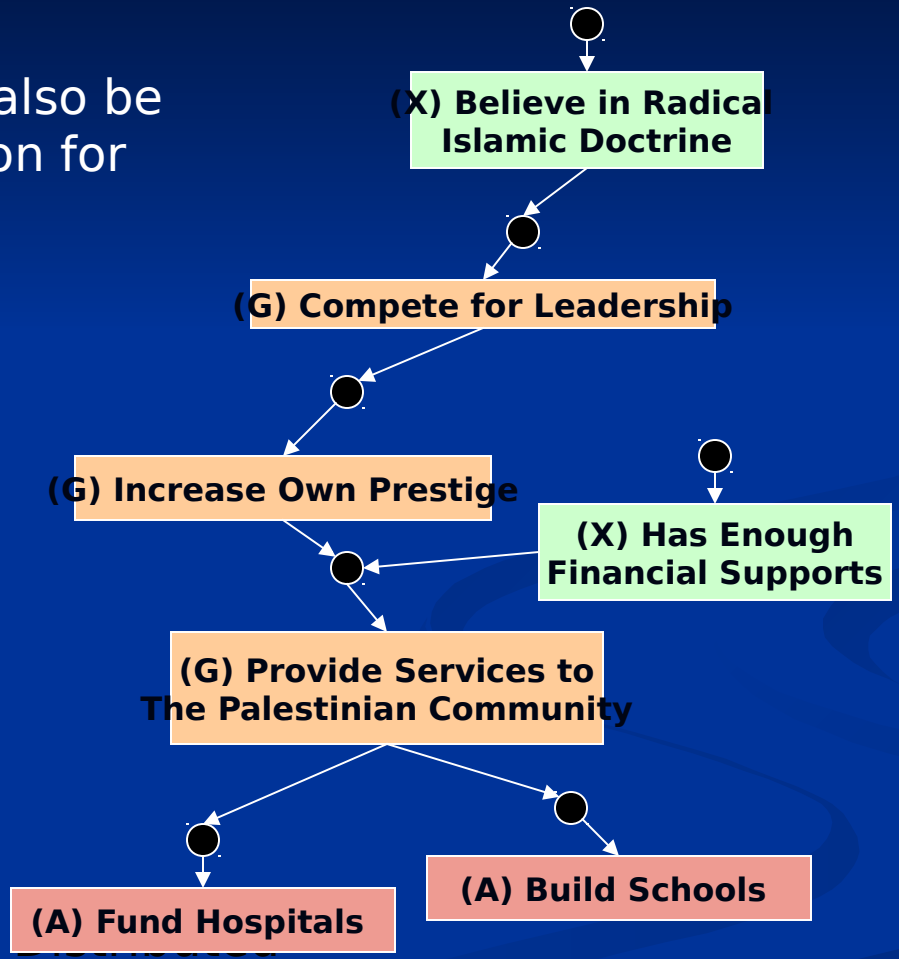

(X) Own Faith in Peace Process (NO)
(X) Believe in Radical Islamic Doctrine (YES)
(B) PA's Authority Questionable (YES)
(B) Israel Willing to Progress Peace Process (NO)
(G) Compete for Leadership (YES)
(B) PA Cooperate with Israel
(G) Damage Peace Process (YES)
(G) Increase Own Prestige
(A) Accuse Peace Deadlock
(G) Damage Trust between Israel and PA (YES)
(G) Damage PA Legitimacy in Palestinian Community (YES)
(A) Accuse PA Corruption
(X) Palestinian Public Support Retaliation Action
(G) Promote Palestinian Civilian Casualty
(B) Israel Overuse Power
(G) Terror Attack against Israel (Yes)
(G) Show Actively Involved In Attacking Israel
(X) Israeli Violence Provoke Doubt on Peace Progress
(A) Terror Attack (YES)
(A) Compete Claiming Responsibility for Terror Attack
(G) Provoke Protest
(A) Suicide Bombing (YES)
(A) Provoke Protest

- PA document: suicide bombing was much more a purely political matter …
- Andrew Kydd and Barbara F. Walter: Violence plays a spoiler role to the peace process. It weakens the moderates (PA) and makes the other side (Israel) become more uncertain.
- James Bennet: Having seen peace initiatives melt before in previous waves of violence, Israelis, like Palestinians, were already deeply skeptical of the new plan.
- Sheikh Ahmed Yassin and Dr. Abdel Aziz Rantisi (Hamas leaders): Suicide bombings were intended to both undermine the legitimacy of the PA and negatively affect the peace process.
- …
- (cited in "Dying to Kill").

One observation: When Palestinian public has hope for the peace process and PA's Authority is unchallengeable, then stop violent action and show cooperation with PA.
In Nov. 1998, 75% Palestinians ceased to support suicide operation;
In 1999, > 70% had faith in the peace process


(G) Increase Own Prestige
(B) PA's Authority Questionable (NO)
(B) PA and Israel Pursue Pease Progress (YES)
(G) Show Cooperating With PA (YES)
(G) Compete for Leadership (NO)
(X) Palestinian Public Has Hope for Peace (YES)
(G) Damage PA Legitimacy in Palestinian Community (NO)
(G) Terror Attack against Israel (NO)
(A) Attend PA Meeting
(A) Terror Attack (NO)
(A) Suicide Bombing (NO)

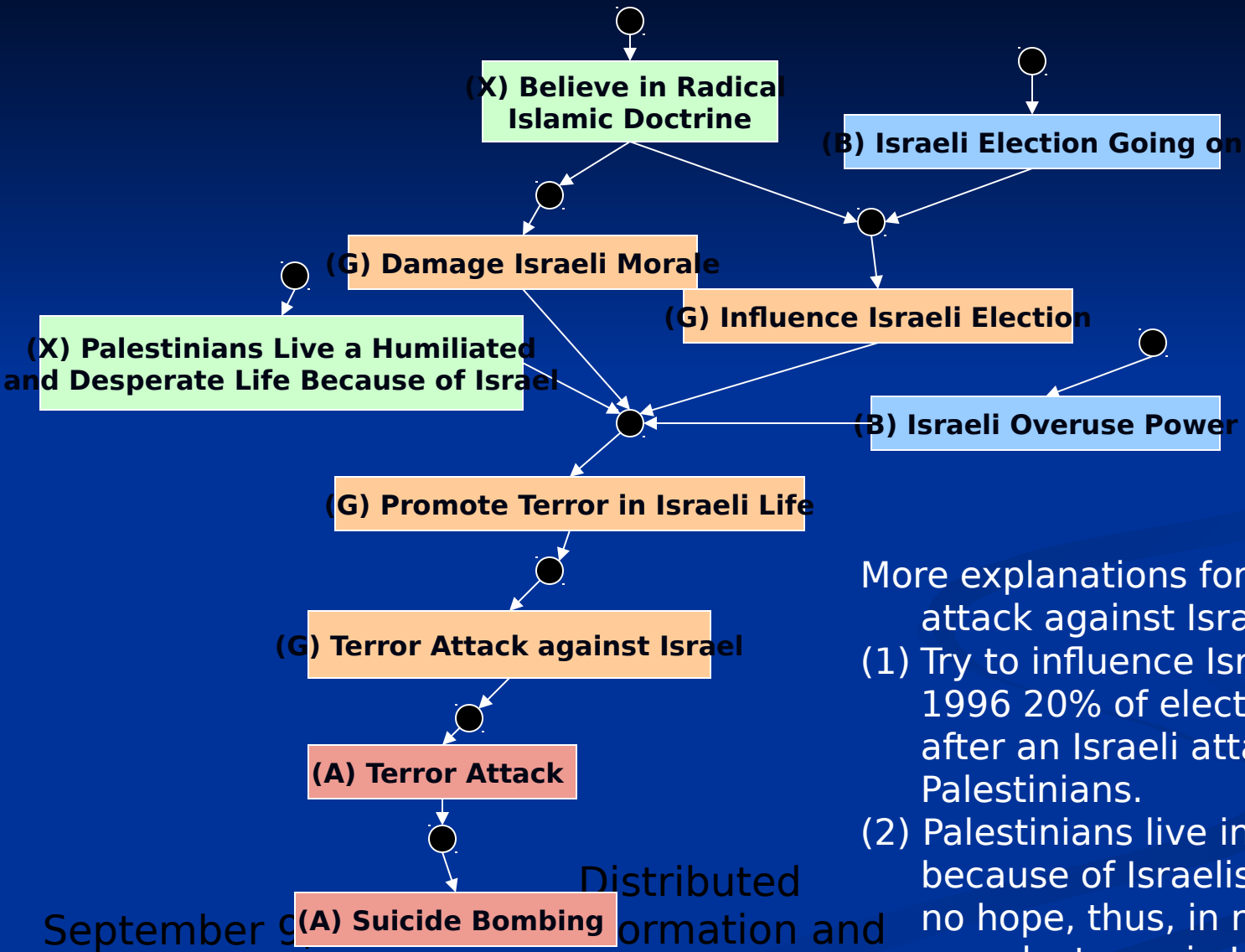
Other actions can also be taken in competition for leadership.


(X) Believe in Radical Islamic Doctrine
(G) Compete for Leadership
(G) Increase Own Prestige
(X) Has Enough Financial Supports
(G) Provide Services to The Palestinian Community
(A) Fund Hospitals
(A) Build Schools

More reasons for using terrorism attacks against Israel: Do not want to take the responsibility of breaking peace progress but try to have Israel start the war.

Richarned Lebow's, "justification of hostility" (cited in "Dying to Kill")

(B) Israeli Overuse Power
(X) Take the Responsibility of Breaking Peace Progress (NO)
(G) Provoke Israel to Start War
(B) Israeli Retaliation
(G) Relate Terror Attack to Israeli Military Action
(G) Terror Attack against Israel
(A) Terror Attack Right After Israeli Military Action
(A) Terror Attack
(A) Suicide Bombing
(A) Suicide Bombing Right After Israeli Military Action

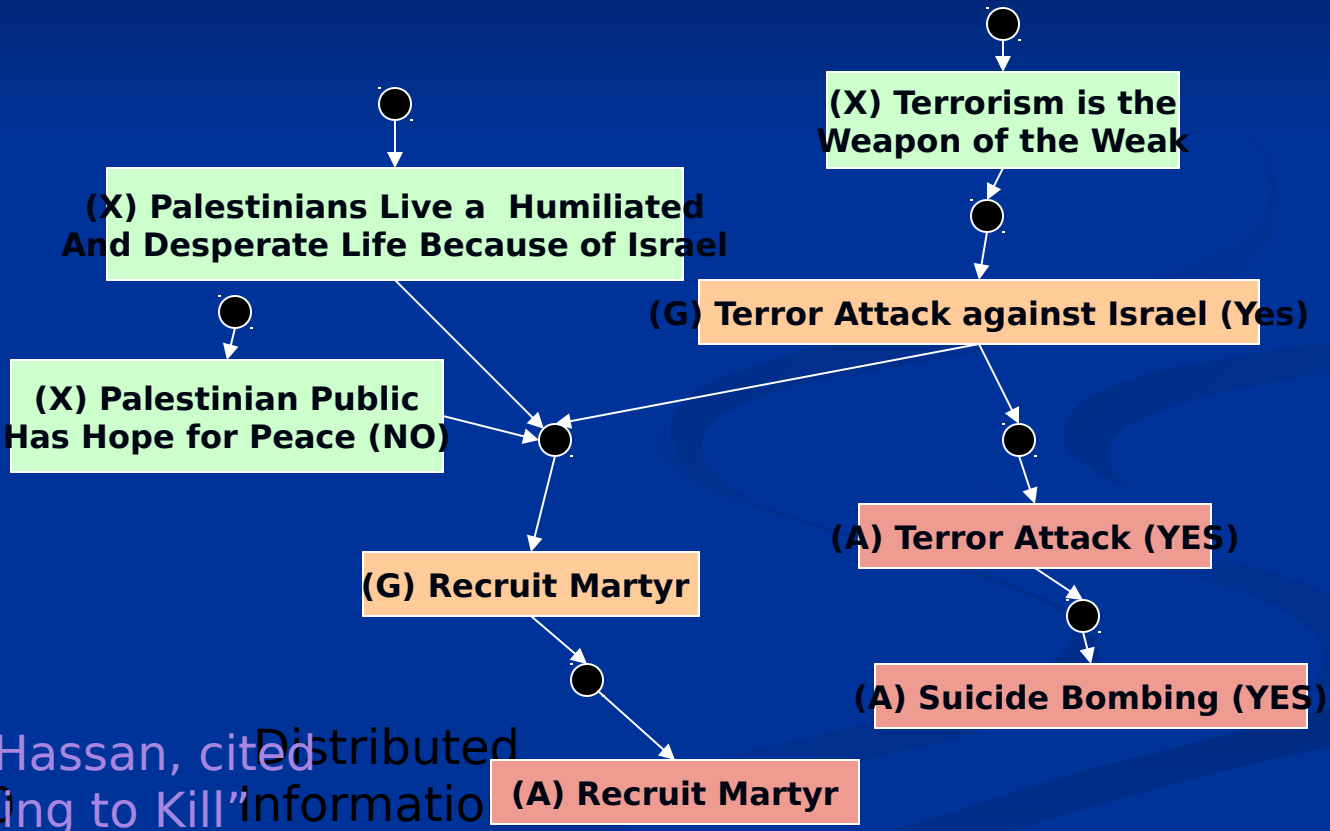

(X) Believe in Radical Islamic Doctrine
(B) Israeli Election Going on
(G) Damage Israeli Morale
(G) Influence Israeli Election
(X) Palestinians Live a Humiliated and Desperate Life Because of Israel
(B) Israeli Overuse Power
(G) Promote Terror in Israeli Life
(G) Terror Attack against Israel
(A) Terror Attack
(A) Suicide Bombing

More explanations for using terrorism attack against Israel:

(1) Try to influence Israeli election; 1996 20% of electorate boycotted after an Israeli attack killed 102 Palestinians.

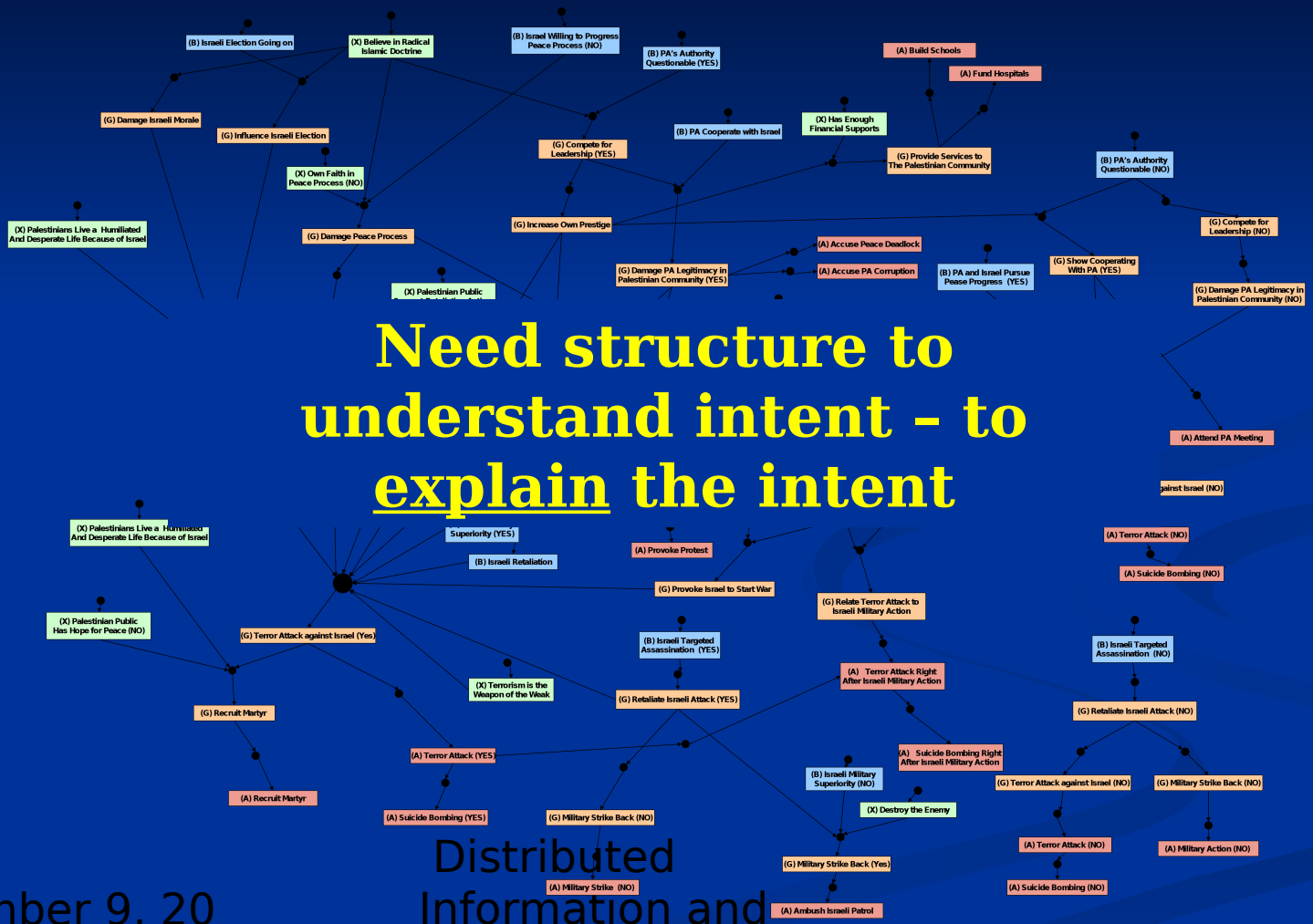
(2) Palestinians live in desperation because of Israelis, and there is no hope, thus, in revenge, want to provoke terror in Israeli life too.

Some factors that influence Palestinian individuals to be recruited as martyrs


(X) Terrorism is the Weapon of the Weak
(X) Palestinians Live a Humiliated And Desperate Life Because of Israel
(G) Terror Attack against Israel (Yes)
(X) Palestinian Public Has Hope for Peace (NO)
(A) Terror Attack (YES)
(G) Recruit Martyr
(A) Suicide Bombing (YES)
(A) Recruit Martyr

Nasra Hassan, cited in "Dying to Kill"

Combined View


(B) Israeli Election Going on
(X) Believe in Radical Islamic Doctrine
(B) Israel Willing to Progress Peace Process (NO)
(B) PA's Authority Questionable (YES)
(A) Build Schools
(A) Fund Hospitals
(G) Damage Israeli Morale
(G) Influence Israeli Election
(X) Own Faith in Peace Process (NO)
(B) PA Cooperate with Israel
(X) Has Enough Financial Supports
(G) Compete for Leadership (YES)
(G) Provide Services to The Palestinian Community
(B) PA's Authority Questionable (NO)
(X) Palestinians Live a Humiliated And Desperate Life Because of Israel
(G) Damage Peace Process
(G) Increase Own Prestige
(A) Accuse Peace Deadlock
(G) Compete for Leadership (NO)
(G) Damage PA Legitimacy in Palestinian Community (YES)
(A) Accuse PA Corruption
(B) PA and Israel Pursue Pease Progress (YES)
(G) Show Cooperating With PA (YES)
(G) Damage PA Legitimacy in Palestinian Community (NO)
(A) Attend PA Meeting
(X) Palestinians Live a Humiliated And Desperate Life Because of Israel
Superiority (YES)
(B) Israeli Retaliation
(A) Provoke Protest
(G) Provoke Israel to Start War
(G) Relate Terror Attack to Israeli Military Action
(A) Terror Attack (NO)
(A) Suicide Bombing (NO)
(X) Palestinian Public Has Hope for Peace (NO)
(G) Terror Attack against Israel (Yes)
(B) Israeli Targeted Assassination (YES)
(B) Israeli Targeted Assassination (NO)
(X) Terrorism is the Weapon of the Weak
(G) Retaliate Israeli Attack (YES)
(A) Terror Attack Right After Israeli Military Action
(G) Retaliate Israeli Attack (NO)
(G) Recruit Martyr
(A) Terror Attack (YES)
(A) Suicide Bombing Right After Israeli Military Action
(A) Recruit Martyr
(A) Suicide Bombing (YES)
(G) Military Strike Back (NO)
(B) Israeli Military Superiority (NO)
(G) Terror Attack against Israel (NO)
(G) Military Strike Back (NO)
(X) Destroy the Enemy
(A) Terror Attack (NO)
(A) Military Action (NO)
(G) Military Strike Back (Yes)
(A) Suicide Bombing (NO)
(A) Military Strike (NO)
(A) Ambush Israeli Patrol

Need structure to understand intent – to <u>explain</u> the intent

Summary

- We initially try to model the terrorist organizations, Hamas and Jihad (PIJ).
- Each network fragment is generated based on one view of what is going on and why it happens this way, such as:
  - Retaliation
  - Competition for leadership
  - Influence Israeli life and election
- The network fragments can be combined/merged together to give a big picture

Summary

- What factors have been discovered thus far:
  - Social: compete for leadership, no hope for peace process
  - Cultural: believe in Islamic doctrine
  - Political: Israeli election
  - Economic: Palestinian's living states
  - Psychological: Humiliation by Israelis
- Ability to take in different models/views
- Not only capture the pattern, but also the reasons

More Challenges

- How to generalize from the specific cases, i.e. identifying potential templates.
- How to set probability values
  - More studies on the empirical data
  - Set values at different levels: low, medium and high,
  - Is the “exact” probability critical?, and
- How to compose network fragments
  - Identify the random variables that have different inputs (parents) in different fragments
  - Group the inputs for such variables

Extract Template from Networks Built in Case Study

- This fragment and the templates obtained from it, contains knowledge:
  - When entity A competes with entity B, there are basically two ways to achieve it: (1) A demonstrates itself to be a better choice; (2) A tries to weaken B's status.
  - In our adversary inferencing model, this represents knowledge that a goal of competing for status can be decomposed into two sub-goals.

Lesson Learned

- Problems in current social science research
  - Lack of empirical data
    - Many articles and books about terrorism since 2001, only 3% contain empirical data
  - Empirical data and analysis typically based on simplistic tools such as linear regression
  - Unstructured data
    - Case studies
  - No general framework on conducting research
    - Many focus on “positive cases” only, which is already biased
    - Non-comparable units of analysis (i.e. time units)
  - Historical changes
- There might be more than one target entity involved
  - In the scenario
    - 1) Organizations, such as Hamas, which we try to model;
    - 2) Individuals, who are the suicide bombers,
- There might be conflicting views for the same cases

Some Empirical Data Suicide Bombing Prediction Model

- From Gupta D. (in press)
  - PIJ suicide bombing at time (t) =

    -3.13 + 0.421* Hamas suicide bombing at time (t-1)

    -1.416* Israeli election + 1.556*political provocation

    +1.582*peace accord
  - Hamas suicide bombing at time (t) =

    -1.157 + 0.75 * PLO shooting at time (t-1)

    + 0.829*election
- What is the appropriate base values at time 0?

Conclusions

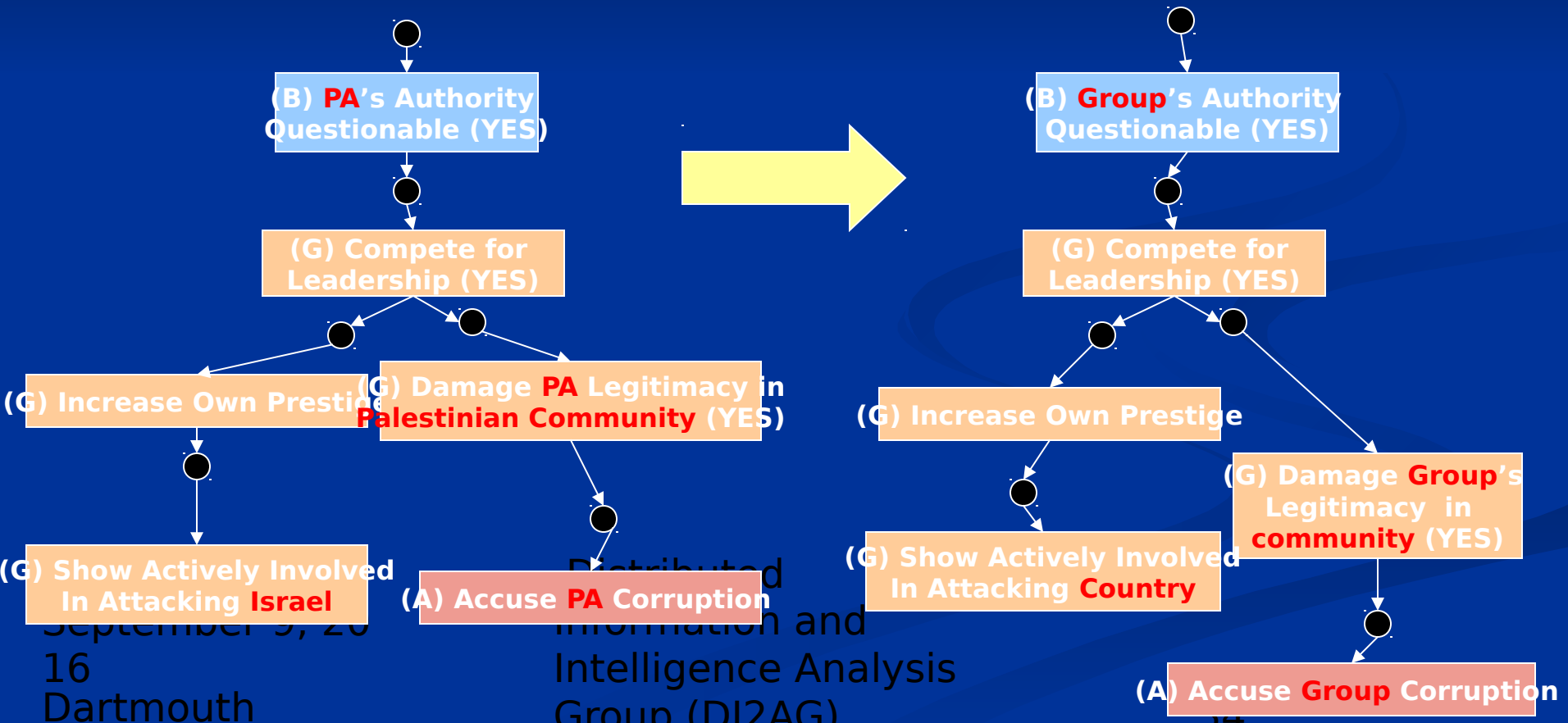
- Continue to develop tools and methodologies for capturing cultural aspects of adversary intent
- Resolve missing data and probabilities by developing models (Bayesian knowledge fragments) that can be evaluated, at least subjectively, by the subject matter experts (social psychologists, politic scientists, etc.)
  - Iterative process
- Continue to overcome vocabulary and even cultural differences between the research disciplines and the researchers themselves

Related Projects

- "Emergent Adverarial Modeling System (EAMS)," AFLR/IF Phase II SBIR with Securboration
- "Dynamic Adversarial Gaming Algorithm (DAGA)," AFOSR Phase I STTR with Securboration
- "Deception Detection in Expert Source Information Through Fusion in Bayesian Knowledge-Base Modelling," AFOSR
- "Fused Intent System," ONR (pending)
- "Intelligence Reporting Inference System (IRIS) Fusion Support Environment," USA RDECOM (pending)

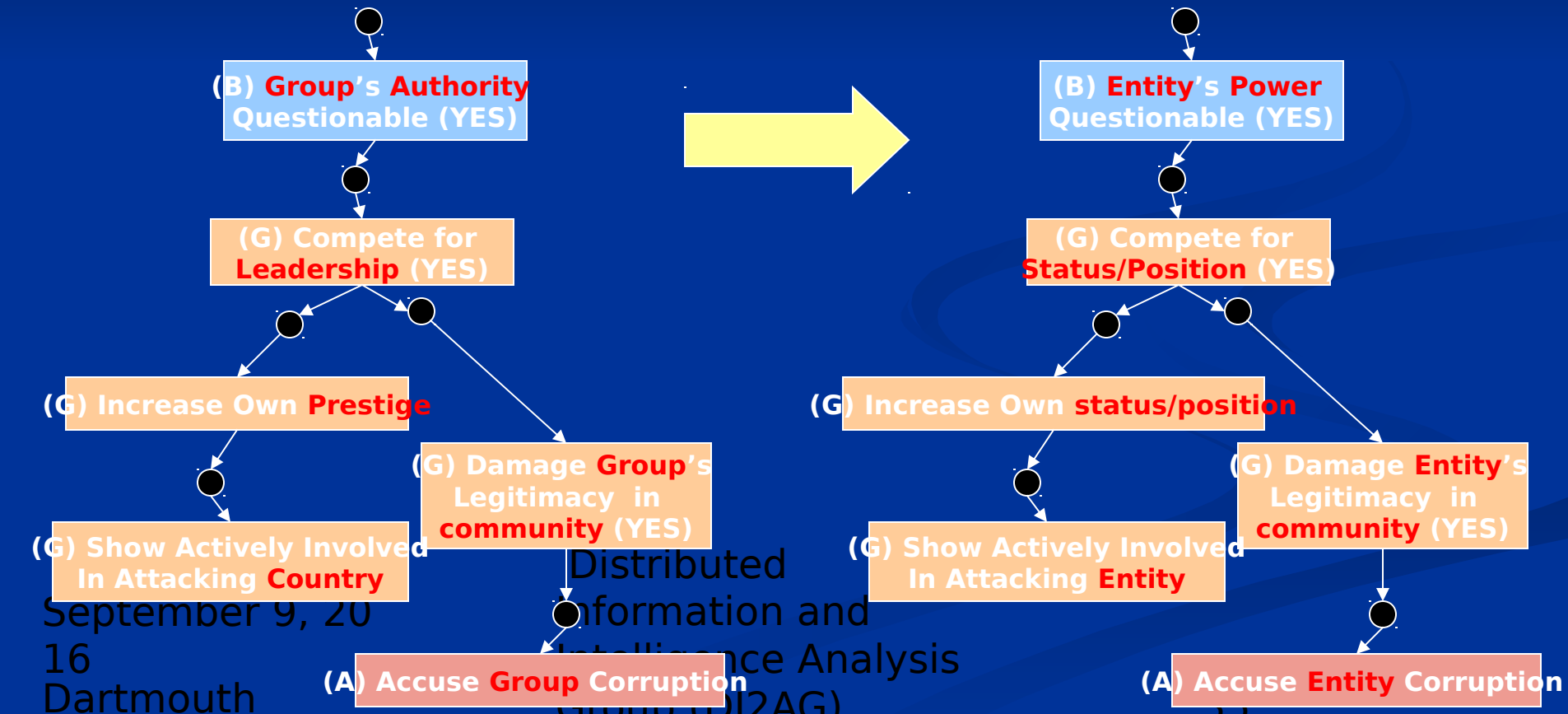
Extract Template from Networks Built in Case Study

Replace specific entities with more general ones, such as PA is an group, Israel is a country, and Palestinian community is a community.


(B) PA's Authority Questionable (YES)
(G) Compete for Leadership (YES)
(G) Increase Own Prestige
(G) Damage PA Legitimacy in Palestinian Community (YES)
(G) Show Actively Involved In Attacking Israel
(A) Accuse PA Corruption
(B) Group's Authority Questionable (YES)
(G) Compete for Leadership (YES)
(G) Increase Own Prestige
(G) Damage Group's Legitimacy in community (YES)
(G) Show Actively Involved In Attacking Country
(A) Accuse Group Corruption

Extract Template from Networks Built in Case Study

The generalization can go further. The templates can then be used in creating more specialized network fragments. Can reflect "flow-down" of group behavior and beliefs to individual behavior.


(B) Group's Authority Questionable (YES)
(G) Compete for Leadership (YES)
(G) Increase Own Prestige
(G) Damage Group's Legitimacy in community (YES)
(G) Show Actively Involved In Attacking Country
(A) Accuse Group Corruption
(B) Entity's Power Questionable (YES)
(G) Compete for Status/Position (YES)
(G) Increase Own status/position
(G) Damage Entity's Legitimacy in community (YES)
(G) Show Actively Involved In Attacking Entity
(A) Accuse Entity Corruption

Extract Template from Networks Built in Case Study

- Which level in the hierarchy is appropriate for generalization/specification?
- When the concept has multiple meanings, which one is the right one? (ambiguity)

Example Hierarchy from WordNet

Israel ➔ administrative district, administrative division, territorial division
➔ country, state, land
➔ district, territory, territorial dominion, dominion
➔ region
➔ location
➔ object
➔ physical entity
➔ entity

Palestinian ➔ Arab, Arabian
➔ Semite
➔ White, white person, Caucasian
➔ person, individual
➔ organism, being
➔ living thing, animate thing
➔ object, physical object
➔causal agent, agency
➔ entity

Some Empirical Data Number of Suicide Bombings

	Hamas	PIJ	Provct n	Peace	Electio n	PLO	Misc
May	0	0	0	0	0	0	0
June	0	0	0	0	0	0	0
Jul-04	0	0	0	0	0	0	0
Aug	0	0	0	0	0	0	0
Sept	0	0	1	0	0	0	0
Oct-04	**1**	0	1	0	0	0	0
Nov	0	0	0	0	0	0	0
Dec-04	0	1	1	0	0	0	0
Jan-05	1	0	1	0	0	0	0
Feb	0	0	1	0	0	0	0
Mar-05	3	0	1	0	0	0	0
Apr-05	2	0	1	0	0	0	0
May-05	1	0	1	0	0	0	0
Jun-05	2	0	1	0	0	0	0
Jul-05	0	1	1	0	0	0	0
Aug-05	3	1	1	0	0	0	0
Sep-05	2	1	1	0	0	0	1
Oct-05	0	1	1	0	0	0	0
Nov-05	0	1	1	0	0	0	0
Dec-05	2	2	1	0	0	0	0

Some Empirical Data Timeline of Significant Events

2001-2002	PFLP conducted 3% of all the attacks,
	significant increase of attacks took place against a political backdrop with few substantive peace negotiations between Israel and PA.
Mar-03	Israeli PM Sharon's incursions into the major West bank town
	Arafat's approve rate stop decline,
	>30% of Palestinian pubic do not trust any current leaders, leave the field wide open for incumbents.
	Support for Fatah might remain, but support for Hamas and Sheikh Yassin increased
Apr-03	Hamas and PFLP claimed credit for Mike's Place on 04/30/2003 (violence became the source of honor)
	p26: Hamas and Fatah get popularity of 22% each (poll)
	<17% Palestinians were optimistic that the violence will end and peaceful negotiations begin again (when?)
05/03-07/03	Hamas called a Hudna toresume operations after attacks on Abdel Aziz Rantisi.
	released PA document - the suicide bombings are a key element in the arena of the struggle between the Israelis and Palestininans

References

- Banks, Sheila B., Stytz, Martin R., Santos, Eugene, Jr., Zurita, Vincent B., and Benslay, James L., Jr., "Achieving Realistic Performance and Decision-Making Capabilities in Computer-Generated Air Forces," *Proceedings of the SPIE 11th Annual International Symposium on Aerospace/Defense Sensing and Controls: AeroSense '97*, Vol. 3085, 195-205, Orlando, FL, 1997.
- Brown, Scott M., Santos, Eugene, Jr., and Bell, Benjamin, "Knowledge Acquisition for Adversary Course of Action Prediction Models," *Proc of the AAAI 2002 Fall Symposium on Intent Inference for Users, Teams, and Adversaries*, Boston, MA, 2002.
- Bell, Benjamin, Santos, Eugene, Jr., and Brown, Scott M., "Making Adversary Decision Modeling Tractable with Intent Inference and Information Fusion," *Proceedings of the 11th Conference on Computer Generated Forces and Behavioral Representation*, 535-542, Orlando, FL, 2002.

References

- Santos, Eugene, Jr., "A Cognitive Architecture for Adversary Intent Inferencing: Knowledge Structure and Computation," *Proceedings of the SPIE: 17th Annual International Symposium on Aerospace/Defense Sensing and Controls: AeroSense 2003*, Vol. 5091, 182-193, Orlando, FL, 2003.
- Surman, Joshua, Hillman, Robert, and Santos, Eugene, Jr., "Adversarial Inferencing for Generating Dynamic Adversary Behavior," *Proceedings of the SPIE: 17th Annual International Symposium on Aerospace/Defense Sensing and Controls: AeroSense 2003*, Vol. 5091, 194-201, Orlando, FL, 2003.
- Santos, Eugene, Jr. and Bell, Benjamin, "Intent Inference for Users, Teams, and Adversaries," *AI Magazine* 24(1), 97-98, AAAI Press, 2003.

References

- Santos, Eugene, Jr. and Negri, Allesandro, "Constructing Adversarial Models for Threat Intent Prediction and Inferencing," *Proceedings of the SPIE Defense & Security Symposium*, Vol. 5423, 77-88, Orlando, FL 2004.
- Santos, Eugene, Jr. and Johnson, Gregory, "Toward Detecting Deception in Intelligent Systems," *Proceedings of the SPIE: Defense & Security Symposium*, Vol. 5423, 131-140, Orlando, FL 2004.
- Revello, Timothy, McCartney, Robert, and Santos, Eugene, Jr., "Multiple Strategy Generation for War Gaming," *Proceedings of the SPIE: Defense & Security Symposium*, Vol. 5423, 232-243, Orlando, FL 2004.
- Lehman, Lynn A., Krause, Lee S., Gilmour, Duane A., Santos, Eugene, Jr., and Zhao, Qunhua "Intent Driven Adversarial Modeling," *Proceedings of the Tenth International Command and Control Research and Technology Symposium: The Future of C2*, McLean, VA, 2005.